MetroXRAINE competition - Motor Imagery dataset

Pasquale Arpaia[1,2], Antonio Esposito[1], and Angela Natalizio[3]

[1] Department of Electrical Engineering and Information Technology (DIETI), University of Naples Federico II, Naples, Italy
[2] Interdepartmental Center for Research on Management and Innovation in Healthcare (CIRMIS), University of Naples Federico II, Naples, Italy
[3] Department of Electronics and Telecommunications (DET), Politecnico di Torino, Turin, Italy

1 Experimental Paradigm

The dataset contains the EEG data of 7 subjects. Five sessions were recorded for each participant. Each session lasted about one hour. A synchronous brain-computer interface (BCI) paradigm was adopted, i.e. the user had to imagine according to an external pace. It consisted of two different motor imagery tasks, namely the imagination of movement of the left hand (class 1), and right hand (class 2). Each session consisted of 6 runs separated by short breaks.

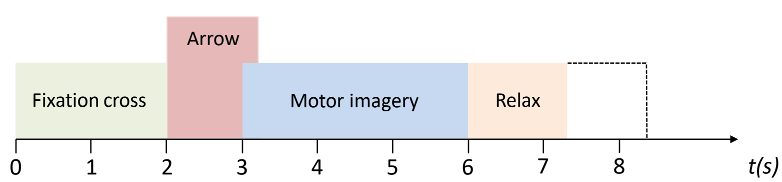
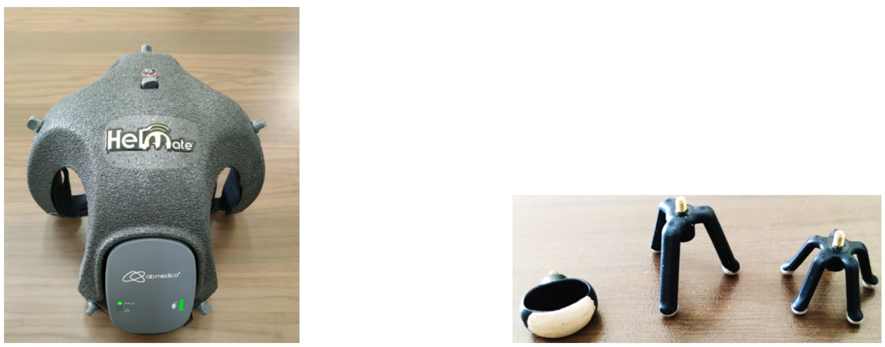


Fig. 1: Timing scheme of the synchronous BCI paradigm.

A single run consisted of 30 randomized trials (15 per class), resulting in a total of 180 trials per session. The participants were sitting in a comfortable chair in front of a computer screen. For each trial, a fixation cross from 0.00 s to 2.00 s, a left or right arrow from 2.00 s to 3.25 s, the textual indication “GO!” up to 6.00 s, and the textual indication “RELAX” disappearing randomly between 7.00 s and 8.00 s were shown sequentially. During the the arrow display, participants were asked to prepare for the motor imagery of the right or left hand depending on the orientation of the arrow. Then, they had to actually imagine the movement when the ”GO!” was shown. The timing of the relax at the end was randomised to avoid any bias between consecutive trials. The timing scheme is illustrated in Fig. 1.

2 Data recording

The EEG signals were recorded through the *ab medica Helmate*. The device is class IIA certified according to the EU regulation 2017/745. The headset



(a) EEG cap

(b) Dry electrodes

Fig. 2: EEG acquisition system.

is equipped with 10 dry electrodes made of conductive-rubber and coated with Ag/AgCl. Channel placement follows the 10/20 International Positioning System and the provided positions are: Fp1, Fp2, Fz, Cz, C3, C4, O1, and O2. AFz is the reference electrode while Fpz is the ground. The electrode placement is shown in Fig. 3.

Electrodes have different shapes to reach the skin passing through the hair. The EEG acquisition system and the dry electrodes are shown in Fig.2a and Fig.2b, respectively.

The Helmate integrates the ADS1298 analog front-end from Texas Instruments, with a multi-channel simultaneous sampling, and a 24-bit $\Delta\Sigma$ analog-to-digital converter (ADC). Analog signals are analog filtered and amplified with a nominal pass-band from 0.2 Hz to 70 Hz with the 50 Hz notch filter enabled. The EEG signals are acquired at a sampling frequency of 512 Sa/s. The device has a rechargeable battery and a Bluetooth connection for data transmission. The Helm8 software manager allows to check the contact impedance between the electrodes and the scalp and the real-time visualization of the EEG data. It also allows a simple pre-processing of the EEG signals.

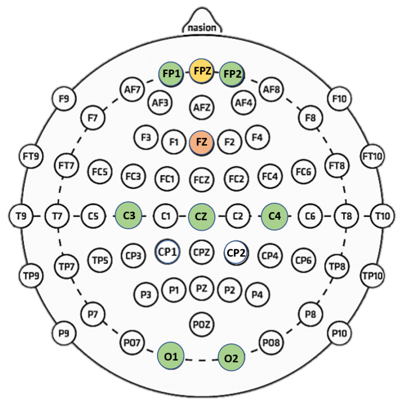


Fig. 3: Electrodes placement

3 Data file description

Data are stored with the *.mat* extension. For each subject, a folder contains four *.mat* files related with the sessions 1 to 4. Once the *.mat* file is loaded, the workspace will contain a cell "data" and a structure "notes" with meta-information. The structure of stored data is shown in Fig. 4.

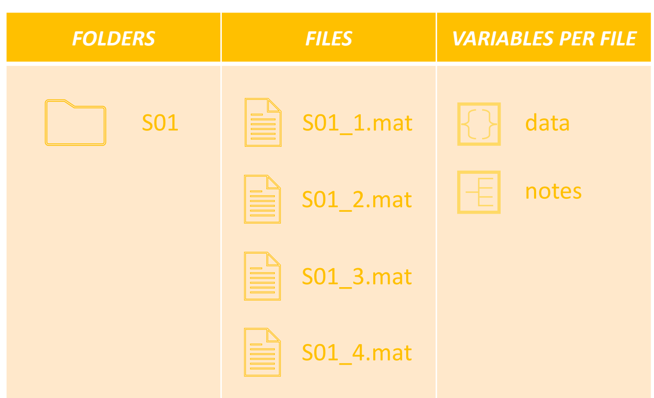


Fig. 4: Structure of stored data.

The data cell contains six structures associated with the six runs acquired for each session. Each structure contains:

– X: the EEG data stream given as a $N \times M$ matrix, where N is the number of channels and M is the number of samples;

- *trial*: a vector of length 30 containing the starting sample per each trial;
- *y*: a vector of length 30 containing the labels per each trial;
- *fs*: the sampling frequency;
- *classes*: a 1×2 cell containing the motor imagery task associated to the two possible labels (0 and 1);
- *artifacts*: a logical vector of length 30. It flags artifact marked by visual inspection. Specifically, 0 corresponds to a clean trial and 1 corresponds to a trial containing an artifact.

Meanwhile, the "notes" structure contains general information like the subject ID, the gender, the age of the subject, the session number, a 1×8 cell containing the name of the channels (sorted according to the rows of the matrix X), reference and ground electrode positions, and the date on which the signals were acquired.

As a final remark: the EEG data of the fifth session for each subject will be only made available during the competition. The challenge will thus involve testing the subject-specific classifier and evaluate the performance on the unknown session.

Cite

If you find useful the EEG data provided by this dataset, please add the following reference to your publications:

Arpaia P., Coyle D., Donnarumma F., Esposito A., Natalizio A., Parvis M., Pesola M., & Vallefuoco E., (2022). Multimodal feedback in assisting a wearable brain-computer interface based on motor imagery. IEEE MetroXRAINE 2022. Accepted